

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

IN RE APPLICATION OF: Masaru SUZUKI, et al.

GAU:

SERIAL NO: New Application

EXAMINER:

FILED: Herewith

FOR: STRUCTURED DATA RETRIEVAL APPARATUS, METHOD, AND PROGRAM

**REQUEST FOR PRIORITY**

COMMISSIONER FOR PATENTS  
ALEXANDRIA, VIRGINIA 22313

SIR:

- ☐ Full benefit of the filing date of U.S. Application Serial Number \_\_\_\_\_, filed \_\_\_\_\_, is claimed pursuant to the provisions of 35 U.S.C. §120.
- ☐ Full benefit of the filing date(s) of U.S. Provisional Application(s) is claimed pursuant to the provisions of 35 U.S.C. §119(e):  
Application No. \_\_\_\_\_ Date Filed \_\_\_\_\_

- ☒ Applicants claim any right to priority from any earlier filed applications to which they may be entitled pursuant to the provisions of 35 U.S.C. §119, as noted below.

In the matter of the above-identified application for patent, notice is hereby given that the applicants claim as priority:

<u>COUNTRY</u>	<u>APPLICATION NUMBER</u>	<u>MONTH/DAY/YEAR</u>
Japan	2002-327127	November 11, 2002

Certified copies of the corresponding Convention Application(s)

- ☒ are submitted herewith
- ☐ will be submitted prior to payment of the Final Fee
- ☐ were filed in prior application Serial No. \_\_\_\_\_ filed \_\_\_\_\_
- ☐ were submitted to the International Bureau in PCT Application Number \_\_\_\_\_  
Receipt of the certified copies by the International Bureau in a timely manner under PCT Rule 17.1(a) has been acknowledged as evidenced by the attached PCT/IB/304.
- ☐ (A) Application Serial No.(s) were filed in prior application Serial No. \_\_\_\_\_ filed \_\_\_\_\_; and
- ☐ (B) Application Serial No.(s) \_\_\_\_\_  
☐ are submitted herewith
- ☐ will be submitted prior to payment of the Final Fee

Respectfully Submitted,

OBLON, SPIVAK, McCLELLAND,  
MAIER & NEUSTADT, P.C.

  
\_\_\_\_\_  
Marvin J. Spivak

Registration No. 24,913

**C. Irvin McClelland**  
**Registration Number 21,124**

Customer Number

**22850**

Tel. (703) 413-3000  
Fax. (703) 413-2220  
(OSMMN 05/03)

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日                      2 0 0 2 年 1 1 月 1 1 日  
Date of Application:

出 願 番 号                      特 願 2 0 0 2 - 3 2 7 1 2 7  
Application Number:

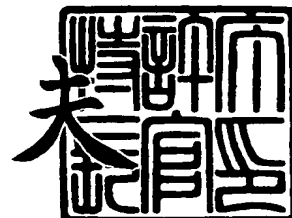
[ST. 10/C]:                      [ J P 2 0 0 2 - 3 2 7 1 2 7 ]

出 願 人                      株式会社東芝  
Applicant(s):

2 0 0 3 年    7 月 1 8 日

特許庁長官  
Commissioner,  
Japan Patent Office

今 井 康 夫



【書類名】 特許願

【整理番号】 A000204019

【提出日】 平成14年11月11日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 7/00

【発明の名称】 構造化データ検索方法、構造化データ検索装置およびプログラム

【請求項の数】 13

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

【氏名】 鈴木 優

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

【氏名】 磯部 庄三

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

【氏名】 布目 光生

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

【氏名】 金輪 拓也

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

【氏名】 小野 顕司

## 【特許出願人】

【識別番号】 000003078

【氏名又は名称】 株式会社 東芝

## 【代理人】

【識別番号】 100058479

## 【弁理士】

【氏名又は名称】 鈴江 武彦

【電話番号】 03-3502-3181

## 【選任した代理人】

【識別番号】 100084618

## 【弁理士】

【氏名又は名称】 村松 貞男

## 【選任した代理人】

【識別番号】 100068814

## 【弁理士】

【氏名又は名称】 坪井 淳

## 【選任した代理人】

【識別番号】 100092196

## 【弁理士】

【氏名又は名称】 橋本 良郎

## 【選任した代理人】

【識別番号】 100091351

## 【弁理士】

【氏名又は名称】 河野 哲

## 【選任した代理人】

【識別番号】 100088683

## 【弁理士】

【氏名又は名称】 中村 誠

## 【選任した代理人】

【識別番号】 100070437

## 【弁理士】

【氏名又は名称】 河井 将次

## 【手数料の表示】

【予納台帳番号】 011567

【納付金額】 21,000円

## 【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 構造化データ検索方法、構造化データ検索装置およびプログラム

【特許請求の範囲】

【請求項 1】

記憶手段に記憶された、複数の構成要素からなる階層化されたデータ構造をもつ複数の構造化データの中から、所望の構造化データを検索するための構造化データ検索方法であって、

前記記憶手段に記憶された各構造化データの構成要素の要素値の種類を推定して、前記構造化データの構成要素と、当該構成要素の要素値に対し推定された種類に対応するラベルとの対応関係を表したテーブルを記憶し、

キーワードと前記ラベルを含む検索要求が入力されたとき、まず、当該検索要求に含まれるラベルに対応する構成要素を前記テーブルから検索し、次に、この検索された構成要素と、前記検索要求に含まれるキーワードを要素値としてもつ構成要素とを含む構造化データを前記記憶手段に記憶された構造化データの中から検索し、

この検索された構造化データのうち、少なくとも、前記検索要求に含まれるラベルに対応する構成要素の要素値を検索結果として出力することを特徴とする構造化データ検索方法。

【請求項 2】

記憶手段に記憶された、複数の構成要素からなる階層化されたデータ構造をもつ複数の構造化データの中から、所望の構造化データを検索するための構造化データ検索方法であって、

前記記憶手段に記憶された各構造化データの構成要素の要素値の種類を推定して、前記構造化データの構成要素と、当該構成要素の要素値に対し推定された種類に対応するラベルとの対応関係を表したテーブルを記憶し、

自然文の検索要求が入力されたとき、当該検索要求としての自然文に含まれる語から前記ラベルを推定するとともに、キーワードに対応する語を推定した後、まず、この推定されたラベルに対応する構成要素を前記テーブルから検索し、次

に、検索された構成要素と、前記推定されたキーワードを要素値としてもつ構成要素とを含む構造化データを前記記憶手段に記憶された構造化データの中から検索し、

この検索された構造化データのうち、少なくとも、前記検索要求から推定されたラベルに対応する構成要素の要素値を検索結果として出力することを特徴とする構造化データ検索方法。

#### 【請求項 3】

前記要素値の種類毎に定められた、当該種類に対応する文字の種類や文字の並びを表した文字列のパターンを、当該種類に対応するラベルに対応つけて予め記憶しておき、

各構造化データの構成要素の要素値の種類を推定する際には、当該構成要素の要素値としての文字列の文字の種類や文字の並びと、前記予め記憶されたパターンと比較して、当該要素値に一致するパターンに対応するラベルを求めることを特徴とする請求項 1 または 2 記載の構造化データ検索方法。

#### 【請求項 4】

前記検索要求としての自然文に含まれる語の中から前記ラベルを推定するために、前記ラベルと当該ラベルを指定していると推定される語とを対応つけて予め記憶しておき、

前記検索要求としての自然文に含まれる語から前記ラベルを推定する際には、当該自然文から前記記憶された語を探索することにより、前記ラベルを求めることを特徴とする請求項 2 記載の構造化データ検索方法。

#### 【請求項 5】

複数の構成要素からなる階層化されたデータ構造をもつ構造化データを蓄積する蓄積手段と、

この蓄積手段に蓄積された各構造化データの構成要素の要素値の種類を推定する第 1 の推定手段と、

前記構造化データの構成要素と、当該構成要素の要素値に対し前記第 1 の推定手段で推定された種類に対応するラベルとの対応関係を表したテーブルを記憶する第 1 の記憶手段と、

キーワードと前記ラベルを含む検索要求を入力する入力手段と、

この入力手段で入力された前記検索要求に含まれるラベルに対応する構成要素を前記テーブルから検索する第1の検索手段と、

この第1の検索手段で検索された構成要素と、前記検索要求に含まれるキーワードを要素値としてもつ構成要素とを含む構造化データを前記蓄積手段で蓄積された構造化データの中から検索する第2の検索手段と、

この第2の検索手段で検索された構造化データのうち、少なくとも、前記検索要求に含まれるラベルに対応する構成要素の要素値を検索結果として出力する出力手段と、

を具備したことを特徴とする構造化データ検索装置。

#### 【請求項6】

複数の構成要素からなる階層化されたデータ構造をもつ構造化データを蓄積する蓄積手段と、

この蓄積手段に蓄積された各構造化データの構成要素の要素値の種類を推定する第1の推定手段と、

前記構造化データの構成要素と、当該構成要素の要素値に対し前記第1の推定手段で推定された種類に対応するラベルとの対応関係を表したテーブルを記憶する第1の記憶手段と、

自然文の検索要求を入力する入力手段と、

この入力手段で入力された検索要求としての自然文に含まれる語から前記ラベルを推定するとともに、キーワードに対応する語を推定する第2の推定手段と、

この第2の推定手段で推定されたラベルに対応する構成要素を前記テーブルから検索する第1の検索手段と、

この第1の検索手段で検索された構成要素と、前記第2の推定手段で推定されたキーワードを要素値としてもつ構成要素とを含む構造化データを前記蓄積手段で蓄積された構造化データの中から検索する第2の検索手段と、

この第2の検索手段で検索された構造化データのうち、少なくとも、前記第2の推定手段で推定されたラベルに対応する構成要素の要素値を検索結果として出力する出力手段と、

を具備したことを特徴とする構造化データ検索装置。

**【請求項 7】**

前記要素値の種類毎に定められた、当該種類に対応する文字の種類や文字の並びを表した文字列のパターンと、当該種類に対応するラベルとを対応つけて記憶する第 2 の記憶手段をさらに具備し、

前記第 1 の推定手段は、各構造化データの構成要素の要素値としての文字列の文字の種類や文字の並びと、前記第 2 の記憶手段に記憶されたパターンとを比較して、当該要素値に一致するパターンに対応するラベルを求めることを特徴とする請求項 5 または 6 記載の構造化データ検索装置。

**【請求項 8】**

前記検索要求としての自然文に含まれる語の中から前記ラベルを推定するために、前記ラベルと当該ラベルを指定していると推定される語とを対応つけて記憶する第 3 の記憶手段をさらに具備し、

前記第 2 の推定手段は、前記検索要求としての自然文から、前記第 3 の記憶手段に記憶された語を探索することにより、前記ラベルを求めることを特徴とする請求項 6 記載の構造化データ検索装置。

**【請求項 9】**

前記第 2 の推定手段で、前記検索要求としての自然文から推定されたラベルとキーワードに対応する語を表示し、これらを修正する手段をさらに具備したことを特徴とする請求項 6 記載の構造化データ検索装置。

**【請求項 10】**

前記第 2 の検索手段で前記構造化データを検索する際に用いる所定の検索言語で記述された検索要求文を表示し、この検索要求文を修正する手段と、

をさらに具備したことを特徴とする請求項 5 または 6 記載の構造化データ検索装置。

**【請求項 11】**

前記第 1 の記憶手段に記憶された前記テーブルを表示する手段をさらに具備したことを特徴とする請求項 10 記載の構造化データ検索装置。

**【請求項 12】**

記憶手段に記憶された、複数の構成要素からなる階層化されたデータ構造をもつ複数の構造化データの中から、所望の構造化データを検索するためプログラムであって、

コンピュータに、

前記記憶手段に記憶された各構造化データの構成要素の要素値の種類を推定するステップと、

前記構造化データの構成要素と、当該構成要素の要素値に対し推定された種類に対応するラベルとの対応関係を記憶するステップと、

キーワードと前記ラベルを含む検索要求を入力するステップと

前記検索要求に含まれるラベルに対応する構成要素を前記対応関係から検索する第1の検索ステップと、

この第1の検索ステップで検索された構成要素と、前記検索要求に含まれるキーワードを要素値としてもつ構成要素とを含む構造化データを前記記憶手段に記憶された構造化データの中から検索する第2の検索ステップと、

この第2の検索ステップで検索された構造化データのうち、少なくとも、前記検索要求に含まれるラベルに対応する構成要素の要素値を検索結果として出力するステップと、

を実行させるプログラム。

### 【請求項13】

記憶手段に記憶された、複数の構成要素からなる階層化されたデータ構造をもつ複数の構造化データの中から、所望の構造化データを検索するためプログラムであって、

コンピュータに、

前記記憶手段に記憶された各構造化データの構成要素の要素値の種類を推定するステップと、

前記構造化データの構成要素と、当該構成要素の要素値に対し推定された種類に対応するラベルとの対応関係を記憶するステップと、

自然文の検索要求を入力するステップと、

前記検索要求としての自然文に含まれる語から前記ラベルを推定するとともに

、キーワードに対応する語を推定する推定ステップと、

この推定ステップで推定されたラベルに対応する構成要素を前記対応関係から検索する第1の検索ステップと、

この第1の検索ステップで検索された構成要素と、前記推定ステップで推定されたキーワードを要素値としてもつ構成要素とを含む構造化データを前記記憶手段に記憶された構造化データの中から検索する第2の検索ステップと、

この第2の検索ステップで検索された構造化データのうち、少なくとも、前記検索要求から推定されたラベルに対応する構成要素の要素値を検索結果として出力するステップと、

を実行させるプログラム。

#### 【発明の詳細な説明】

##### 【0001】

##### 【発明の属する技術分野】

本発明は、複数の構成要素からなる階層化されたデータ構造（論理構造）を持つ構造化データを検索する構造化データ検索装置に関し、特に、構成要素の要素名などのデータ構造が統一されていない様々なデータ構造の構造化データの中から所望の構造化データを検索するための構造化データ検索装置に関する。

##### 【0002】

##### 【従来の技術】

従来のリレーショナルデータベース管理システム（RDBMS）ではデータベースの構造が予め定められており、利用者はデータの構造を利用した検索を行なうことができる。例えばある商品データベースにおいて、商品価格を保持するための属性として「PRICE」という属性が用意されていれば、データ作成者は商品価格を「PRICE」属性に記録し、検索利用者は商品価格を参照するために「PRICE」属性を参照すればよい、ということが予め決められている。このような検索を行なうための検索言語としては構造化問い合わせ言語（SQL）がある。また、検索利用者がSQLを記述するのではなく、検索利用者は自然言語で検索要求を入力し、システムが検索要求を解釈してSQLに変換する、という技術もある（例えば、特許文献1参照）。この場合も、データベースの構造が

既知であるので、自然言語の解釈結果をデータベースの構造に合わせるための知識を予め用意しておくことができる。また、近年注目されている拡張可能なマークアップ言語（XML: extensible markup language）を利用したデータベース管理システムではXQUERYなどの検索言語が用意されている。

#### 【0003】

複数の構成要素からなる階層化された論理構造をもつ構造化データの1つであるXMLデータでは、構成要素の要素名などのデータ構造は必ずしも予め定められている必要はなく、データ作成者が独自に要素名などのデータ構造を定める（あるいは拡張する）ことができる。XMLデータでは、上記属性を構成要素の要素名、すなわち、タグとして定義することができる。例えば商品データベースにおいて、商品価格を保持するためのタグは、＜PRICE＞タグでも、＜価格＞タグでもよく、さらに、＜TAG1＞であってもよい。このように、タグ名は、自由に設定できるため、＜TAG1＞など必ずしも内容を表さないタグ名が利用される場合もある。

#### 【0004】

このため、検索利用者には商品価格がどのタグに記述されているかわからない。また、データを一瞥しただけではデータの構造を知ることは困難である。

#### 【0005】

このように、従来は、構造化データを検索する際には、検索しようとするユーザは、予め検索したい構造化データの構成要素の要素名などのデータ構造を知る必要があった。ユーザが知り得るデータ構造の範囲内に構造化データの検索を制限することは、XMLデータを利用するメリットが大きく損われるというものである。

#### 【0006】

##### 【特許文献1】

特開平5-54078号公報

#### 【0007】

##### 【発明が解決しようとする課題】

従来の検索手法では、異なるデータ構造の複数の構造化データの中から、所望

の構造化データを検索する際には、ユーザは、この異なるデータ構造を全て知る必要があり、従って、所望のデータを要素値としてもつ構成要素を含む構造化データを容易に、しかも確実に検索することが困難であるという問題点があった。

#### 【0008】

そこで、本発明は上記問題点に鑑み、異なるデータ構造の複数の構造化データの中から、所望の構造化データを検索する際には、そのデータ構造に依存することなく、所望の情報データを要素値としてもつ構成要素を含む構造化データを容易に、しかも確実に検索することができる構造化データ検索方法およびそれを用いた構造化データ検索装置を提供することを目的とする。

#### 【0009】

##### 【課題を解決するための手段】

(1) 本発明は、記憶手段に記憶された、複数の構成要素からなる階層化されたデータ構造をもつ複数の構造化データの中から、所望の構造化データを検索するためのものであって、前記記憶手段に記憶された各構造化データの構成要素の要素値の種類を推定して、前記構造化データの構成要素と、当該構成要素の要素値に対し推定された種類に対応するラベルとの対応関係を表したテーブルを記憶し、キーワードと前記ラベルを含む検索要求が入力されたとき、まず、当該検索要求に含まれるラベルに対応する構成要素を前記テーブルから検索し、次に、この検索された構成要素と、前記検索要求に含まれるキーワードを要素値としてもつ構成要素とを含む構造化データを前記記憶手段に記憶された構造化データの中から検索し、この検索された構造化データのうち、少なくとも、前記検索要求に含まれるラベルに対応する構成要素の要素値を検索結果として出力することを特徴とする。

#### 【0010】

本発明によれば、ユーザは、所望のラベルとキーワードを含む検索要求を入力すれば、当該ラベルに対応する構成要素と当該キーワードを含む構造化データを検索することができる。すなわち、本発明によれば、異なるデータ構造の複数の構造化データの中から、所望の構造化データを検索する際には、そのデータ構造に依存することなく（正確な要素名を知る必要なく、検索したいデータの種類の

対応するラベルを指定すれば)、所望の情報データを要素値としてもつ構成要素を含む構造化データを容易に、しかも確実に検索することができる。

#### 【0011】

好ましくは、前記要素値の種類毎に定められた、当該種類に対応する文字の種類や文字の並びを表した文字列のパターンを、当該種類に対応するラベルに対応つけて予め記憶しておき、各構造化データの構成要素の要素値の種類を推定する際には、当該構成要素の要素値としての文字列の文字の種類や文字の並びと、前記予め記憶されたパターンとを比較して、当該要素値に一致するパターンに対応するラベルを求める。

#### 【0012】

(2) 本発明は、記憶手段に記憶された、複数の構成要素からなる階層化されたデータ構造をもつ複数の構造化データの中から、所望の構造化データを検索するためのものであって、前記記憶手段に記憶された各構造化データの構成要素の要素値の種類を推定して、前記構造化データの構成要素と、当該構成要素の要素値に対し推定された種類に対応するラベルとの対応関係を表したテーブルを記憶し、自然文の検索要求が入力されたとき、当該検索要求としての自然文に含まれる語から前記ラベルを推定するとともに、キーワードに対応する語を推定した後、まず、この推定されたラベルに対応する構成要素を前記テーブルから検索し、次に、検索された構成要素と、前記推定されたキーワードを要素値としてもつ構成要素とを含む構造化データを前記記憶手段に記憶された構造化データの中から検索し、この検索された構造化データのうち、少なくとも、前記検索要求から推定されたラベルに対応する構成要素の要素値を検索結果として出力することを特徴とする。

#### 【0013】

本発明によれば、ユーザは、どのような構造化データを検索したいのかを自然文で表現した検索要求を入力すれば、実際の検索に用いるためのラベルとキーワードとを当該自然文から推定する。そして、このラベルに対応する構成要素と当該キーワードを含む構造化データを検索する。すなわち、本発明によれば、異なるデータ構造の複数の構造化データの中から、所望の構造化データを検索する際

には、そのデータ構造に依存することなく（正確な要素名を知る必要なく、検索したいデータの種類やキーワードなどを表した自然文を入力すれば）、所望の情報データを要素値としてもつ構成要素を含む構造化データを容易に、しかも確実に検索することができる。

#### 【0014】

好ましくは、前記要素値の種類毎に定められた、当該種類に対応する文字の種類や文字の並びを表した文字列のパターンを、当該種類に対応するラベルに対応つけて予め記憶しておき、各構造化データの構成要素の要素値の種類を推定する際には、当該構成要素の要素値としての文字列の文字の種類や文字の並びと、前記予め記憶されたパターンとを比較して、当該要素値に一致するパターンに対応するラベルを求める。

#### 【0015】

また、好ましくは、前記検索要求としての自然文に含まれる語の中から前記ラベルを推定するために、前記ラベルと当該ラベルを指定していると推定される語とを対応つけて予め記憶しておき、前記検索要求としての自然文に含まれる語から前記ラベルを推定する際には、当該自然文から前記記憶された語を探索する。

#### 【0016】

##### 【発明の実施の形態】

以下、本発明の実施の形態について、図面を参照して説明する。

#### 【0017】

##### （第1の実施形態）

第1の実施形態では、構造化データの記述言語としてXMLを用い、構造化データの具体例として複数の店舗が扱う種々の商品の情報を検索する、商品情報検索装置を考える。このような商品情報（構造化データ）は、例えばインターネット上に各店舗が公開しているカタログデータとして収集することができる。このように集められた商品情報は、各店舗毎に構造が異なり、さらにはそれらの構造が予告なく更新されることが考えられる。

#### 【0018】

図1は、第1の実施形態に係る商品情報検索装置の構成例を示したものである

。

#### 【0019】

構造化データ蓄積部101は、ここでは明示しないデータ収集手段によってインターネット上から収集された、XMLで記述された商品情報データを蓄積する。商品情報データは、複数の構成要素からなる階層化されたデータ構造をもつ構造化データである。各構成要素は、当該構成要素の名称に対応する要素名（タグあるいはタグ名）をもち、要素値として、データ（テキストデータや画像データや音声データ）をもつものもある。

#### 【0020】

なお、以下の説明において、各構成要素をタグ名あるいはタグで区別して呼ぶことがある。

#### 【0021】

第1の推定知識保持部102は、商品情報データ中の要素値の種類を推定するための推定知識を保持する。

#### 【0022】

第1の推定部103は、第1の推定知識保持部102が保持する推定知識を用いて、構造化データ蓄積部101に蓄積された商品情報データの構成要素の要素値の種類を推定し、その推定結果を推定結果記憶部104に記憶する。推定結果記憶部104には、商品情報データの構成要素と、当該構成要素の要旨に対し第1の推定部103で推定された種類に対応するラベルとの対応関係を表したテーブルが記憶されている。

#### 【0023】

検索要求入力部105は、ユーザからの検索要求の入力を受け付け、検索要求変換部107に送る。

#### 【0024】

変換知識保持部106は、検索要求を検索部108が解釈できる検索言語に変換するための変換知識を保持する。ここでは、検索言語として具体的にはXQUERY（URL「<http://www.w3.org/XML/Query>」、<http://www.w3.org/TR/xquery>」）を考える。

## 【0025】

検索要求変換部107は、変換知識保持部106が保持する変換知識を参照して検索要求を検索部108が解釈できる検索言語に変換する。

## 【0026】

検索部108は、検索要求変換部107の変換結果である、XQUERYで記述された検索命令に基づき、構造化データ蓄積部101に蓄積された商品情報データや、後述する推定結果記憶部104に記憶された推定結果などの構造化データを検索する。

## 【0027】

検索部108は、公知技術のXMLデータベース管理システムを用いてもよく、ここでは、検索手法そのものは、本願発明の要旨ではないので、説明は省略する。

## 【0028】

検索結果出力部109は、検索部108での検索結果をユーザに提示するためのものである。

## 【0029】

図2は、構造化データ蓄積部101に蓄積された、構造化データの一例として、商品情報データの例を図2(a)～(d)のそれぞれに示したものである。

## 【0030】

商品情報データは、例えば、図2(a)に示すように、XMLで表現され、収集手段により、各店舗から収集された生データ211と、生データを店舗毎に整理するために、例えば収集手段によって付与された、上記生データ以外の整理用データ201が含まれている。生データ211は、<データ>タグで囲まれた部分である。図2(b)～(d)についても図2(a)と同様である。

## 【0031】

構造化データ蓄積部101には、上記のような複数の構造化データが記憶されている。

## 【0032】

図3は、第1の推定知識保持部102に保持されている、商品情報データ中の

要素値の種類を推定するための推定知識の例を示したものである。

#### 【0033】

ここでは推定知識をテーブル形式で示しているが、この場合に限らず、推定知識は、例えばXMLを用いて表現されたXMLデータとして、構造化データ蓄積部101と同様にして保持することができる。

#### 【0034】

図3に示す推定知識は、例えばPerl言語（「Perlプログラミング」Larry Wall et al., ソフトバンク株式会社、pp. 31-32 参照）を用いて表現されたパターンと、当該パターンに対応するラベルの対になっている。パターンとは、要素値として表されている内容の意味や役割などから分類することのできる種類を、定型化して表したものである。1つの種類に対し1つのラベルが対応し、1つのラベルに対し、1または複数のパターンが対応付けられている。

#### 【0035】

例えば、推定知識301は、「価格」というラベルと、当該ラベル（すなわち、種類）に対応するパターンを対応付けたもので、このパターンは、例えば、「1000円」や「1,000円」のように、「1つ以上の数字またはカンマ（,）の列の直後に「円」という文字」という文字列を表したものである。また、推定知識302は、上記推定知識301と同様、「価格」というラベルと、当該ラベル（すなわち、種類）に対応する他のパターンを対応付けたもので、このパターンでは、例えば、「¥1000」や「¥1,000」のように、上記推定知識301で表されたパターンの先頭に、さらに「¥」のある文字列を表している。

#### 【0036】

推定知識303では、「時間」というラベルと、当該ラベルに対応するパターンを対応付けたもので、このパターンは、例えば、「3.5時間」のような文字列を表したものである。また、推定知識304では、「時間」というラベルと、当該ラベルに対応する他のパターンを対応付けたもので、このパターンは、例えば、「3時間5分」のような文字列を表したものである。

#### 【0037】

推定知識305は、「長さ」というラベルと、当該ラベルに対応するパターンを

対応付けたもので、このボタンは、例えば、「10.5cm」「10.2mm」「10.1m」のような文字列を表したものである。

#### 【0038】

推定知識306は、「容量」というラベルと、当該ラベルに対応するボタンを対応付けたもので、このボタンは、例えば、「10GB」「11MB」のような文字列を表したものである。

#### 【0039】

推定知識307は、「周波数」というラベルと、当該ラベルに対応するボタンを対応付けたもので、このボタンは、例えば、「1.8GHz」「1.9MHz」のような文字列を表したものである。

#### 【0040】

このように、図3に示す推定知識は、それぞれ、要素値が文字列の場合に、当該文字列を構成する文字の種類や文字の並びなどをボタン化し、それを要素値の種類に（ラベル）に対応付けて構成されたものである。

#### 【0041】

次に、図4を参照して、第1の推定知識保持部102に記憶されている、上記推定知識を用いた第1の推定部103の処理動作について説明する。

#### 【0042】

まず、第1の推定部103は構造化データ蓄積部101からひとつずつ構造化データとしての商品情報データを取り出す（ステップS1）。ここで構造化データ蓄積部101に蓄積された商品情報データの例として図2（a）～（e）を考えると、このステップS1では、XMLデータベースとしての構造化データ蓄積部101から、順に＜商品情報＞タグ以下のサブツリー（すなわち、＜商品情報＞タグで囲まれた部分）を取り出す処理になる。

#### 【0043】

次に、ステップS1で取得した＜商品情報＞タグ以下のサブツリーから、店舗名とデータを取り出す（ステップS2）。すなわち、ここでは、＜店舗名＞タグの値（要素値）と、＜データ＞タグ以下のサブツリー（すなわち、＜データ＞タグで囲まれた部分）を取り出す処理になる。

## 【0044】

ステップS3では、ステップS2で取り出した<データ>タグ以下のサブツリーの中から、順に、構成要素（タグ）とその値を取り出す。なお、ここでは、1つの構成要素が他の構成要素を包含している場合（構成要素が階層構造を有する場合）は、当該包含されているタグを出現順に取り出し、その値として含まれる子タグ以下を取り除くこととする。

## 【0045】

ステップS4で、第1の推定知識保持部102から推定知識を1つずつ順に取り出し、ステップS5では、第1の推定部103において、ステップS3で取得したタグの値（要素値）がステップS4で取り出した推定知識のパターンとマッチするかチェックする。

## 【0046】

ステップS5でのチェックの結果、要素値と推定知識のパターンとがマッチした場合、ステップS6に進み、当該要素値のタグ名と当該推定知識に対応するラベルとの対応関係を推定結果記憶部104に記憶するとともに、ここでは、例えば当該対応関係に、当該タグ名をもつ商品情報に対応する店舗名（この店舗名は、当該商品情報中の<店舗名>タグの値として記述されている）を対応付けて記憶する。その後、ステップS7へ進む。

## 【0047】

なお、ステップS6では、推定結果記憶部104に、当該要素値のタグ名と当該推定知識に対応するラベルと当該タグ名をもつ商品情報に対応する店舗名との対応関係を記憶するようになっているが、この場合に限らず、ここでは、少なくとも、どの商品情報（どの店舗名の商品情報）の構成要素かがわかるように対応関係が記憶されていればよい。例えば、構造化データ中の各構成要素を当該構造化データの階層構造（ツリー構造）におけるノードとみなし、当該構造化データ中の目的の構成要素の位置を、当該ツリー構造の先頭から目的の構成要素のノードに至るまでの経路上の構成要素を並べて表現することもできる。これをパスと呼ぶ。このパスを用いて上記ラベルと対応つける構成要素（タグ名）を表現してもよい。

## 【0048】

推定結果記憶部104に記憶される上記対応関係の記憶例を図5に示す。なお、図5、図6の推定結果の説明は後述する。

## 【0049】

ステップS5で要素値が推定知識のパターンとマッチしなかった場合には、ステップS7へ進む。ステップS7では、ステップS3で取り出した要素値に対し、第1の推定知識保持部102に記憶されている全ての推定知識を用いてチェックし終わるまで、ステップS4～ステップS6を繰り返すようになっている。

## 【0050】

ステップS7において、ステップS3で取り出された要素値に対し、全ての推定知識を用いたチェックが終了したときは、ステップS8に進む。

## 【0051】

ステップS8では、ステップS2で取り出した<データ>タグ以下のサブツリーから全てのタグを取り出したか判断し、まだ取り出していないタグがある場合にはステップS3に戻るようになっている。

## 【0052】

ステップS8で全てのタグについての処理が終了したと判断された場合、ステップS9に進み、構造化データ蓄積部101から全ての商品情報を取り出したかを判断する。まだ取り出していない商品情報がある場合にはステップS1に戻り、上記ステップS1～ステップS8を繰り返し、全ての商品情報が処理された場合には終了する。

## 【0053】

ここでは、構造化データ蓄積部101に予め蓄積された構造化データ（商品情報）を全て取り出して処理するフローを説明したが、構造化データ蓄積部101に新たな商品情報が追加された場合に、当該追加された商品情報について順次、図4と同様の手順で処理すればよい。

## 【0054】

次に、図4に示した、第1の推定部103の処理動作を、図2に示した商品情報と、図3に示した推定知識を用いて具体的に説明する。

## 【0055】

ステップS1で構造化データ蓄積部101から商品情報がひとつ取り出される。ここでは図2(a)が取り出された場合を考える。

## 【0056】

ステップS2で、図2(a)から店舗名として「AAデンキ」、データとして、<データ>タグで囲まれた部分211が取り出される。

## 【0057】

ステップS3で、<データ>タグで囲まれた部分211内から順にタグとその値が取り出される。ここでは最初に<商品名>タグとその値「PC-A100」が取り出された場合を考える。

## 【0058】

ステップS4では、図3に示したような推定知識の中から上から順に推定知識が取り出される。ここでは最初に推定知識301が取り出された場合を考えるが、この場合、ステップS5では、当該推定知識301のパターンと「PC-A100」とはマッチしないためステップS7に進む。

## 【0059】

その後、ステップS4～ステップS7が繰り返されるが、「PC-A100」は図3の推定知識のいずれともマッチせず、ステップS8へと進む。

## 【0060】

ステップS8からステップS3に戻り、次に、<データ>タグで囲まれた部分211から、<販売価格>タグとその値「123, 000円」が取り出される。

## 【0061】

ステップS4でもう一度推定知識が順に取り出される。ここでは先程と同様に推定知識301が取り出された場合を考える。

## 【0062】

推定知識301が「123, 000円」とマッチするため、次に、ステップS6に進む。

## 【0063】

ステップS6では、当該<販売価格>タグのXQUERY (XPATH) 表現

(パス表現の一種)である「／商品情報[店舗名＝“AAデンキ”]／データ／販売価格」と、推定知識301のラベルである「価格」を対として、推定結果記憶部104に記憶し、ステップS7に進む。

#### 【0064】

以下同様に、図3の推定知識が順に全て取り出されるまで、ステップS4～ステップS7を繰り返し、ステップS7で推定知識が全て取り出される。また、＜データ＞タグで囲まれた部分211から全てのタグが取り出されるまで、ステップS3～ステップS108が繰り返される。さらに、構造化データ蓄積部101に蓄積されている、例えば、図2(b)～(d)に示す全ての商品情報に対し、ステップS1～ステップS9を繰り返す。

#### 【0065】

以上のラベル推定結果は、推定結果記憶部104に、例えば、図5に示すように記憶される。ここで、例えば図5における最初の推定結果(対応関係)501では、その店舗名から、図2(a)及び(b)に示した商品情報の両方に適用されるが、実際に、図4の推定処理動作によりラベルの対応付けられたものは、図2(a)のみである。このことは、推定結果を図5のように表現することで、実際に推定知識のパターンに要素値がマッチしないタグであっても、当該要素値の種類を推定してラベルを対応付けることができることを示している。

#### 【0066】

図5は、推定結果を模式的に示しているが、実施方法としては例えば図6に示したように、推定結果をXMLによって表現したXMLデータとして、構造化データ蓄積部101と同様にして、推定結果記憶部104に記憶することができる。この場合、検索部108は、構造化データ蓄積部101に記憶されている構造化データを検索するのと同様にして、推定結果記憶部104から推定結果を検索することができる。本実施形態の以後の説明では、推定結果は図6のようにXMLで表現され、構造化データ蓄積部101と同じデータベース上に記録されているものとする。

#### 【0067】

なお、図4では、ステップS5で推定知識のパターンがマッチすると直ちにステ

ップ S 6 へ進み、推定結果を記憶すると説明したが、この場合に限らない。例えば、同じタグに異なる推定知識のラベルに対応するボタンがマッチした場合に、ステップ S 2 で取り出した店舗名毎にタグとラベルの対応関係を集計し、統計的に（例えば一番多くマッチしたラベルを選択するなどの方法で）推定を行なってもよい。

#### 【0068】

図 7 は、検索要求入力部 105 に検索要求を入力するための手順を説明するための図である。検索要求を入力するために、例えば、図 7 (a) に示すような検索要求入力画面が検索要求入力部 105 により表示される。この検索要求入力画面は、キーワード入力領域と選択式の GUI によってラベルを入力することにより、検索要求の入力を受け付ける例である。

#### 【0069】

図 7 (a) は、検索要求入力画面の初期状態を示している。ここで利用者が、「「PC-B200」の「価格」を知りたい」場合、キーワード入力領域 601 にキーワードとして「PC-B200」と入力し、続いて、ラベル入力領域 602 にラベルを入力するために、まず、当該ラベル入力領域 602 に設けられたボタン 603 を選択する。すると、図 7 (b) に示すような選択可能なラベルの一覧が表示されるので、利用者は、この一覧の中から所望のラベル（例えば、ここでは「価格」）を選択する。

#### 【0070】

このようにして、入力されたキーワードとラベルとを含む検索要求（図 7 (c) 参照）は、内部では例えば「ラベル＝“価格”、キーワード＝“PC-B200”」といった表現で保持することができる。もちろん XML 形式で保持してもよい。

#### 【0071】

図 8、図 9 は変換知識保持部 106 に保持された変換知識の例を示したものである。

#### 【0072】

図 8 は、予備検索のための第 1 の検索命令を作成するために用いる変換知識（

第1の変換知識)であり、図9は、構造化データ蓄積部101に蓄積された商品情報を検索するための第2の検索命令を作成するために用いる変換知識(第2の変換知識)を示している。

#### 【0073】

なお、本実施形態では、利用者により入力された検索要求に対応する検索は、第1の検索命令を用いた予備検索(第1の検索)と、第2の検索命令を用いた第2の検索との2回行われる。

#### 【0074】

図8に示す第1の変換知識は、推定結果記憶部104から、利用者からの検索要求に含まれていたラベルと対応関係にある構成要素(タグ名)と店舗名を検索するための第1の検索命令を生成するための変換知識である。

#### 【0075】

図9に示す第2の変換知識は、予備検索の結果を利用して利用者からの検索要求を検索命令(第2の検索命令)に変換するための変換知識である。各変換知識は、ラベル毎に対応付けられている。

#### 【0076】

なお、この実施形態では、第1の変換知識及び第2の変換知識は、XQUERYのFLWR構文と呼ばれる形式の一部を置換文字列とした形式で表現されている。

#### 【0077】

図8に示した第1の変換知識では、「##ROLE##」と表現されている部分が、ユーザにより検索要求として入力されたラベルが代入される、置換文字列である。後述する手順でこれらの置換文字列が、上記検索要求に含まれるラベルに置換され、例えば、ここでは、検索言語XQUERYで記述された第1の検索命令が生成される。なお、変換知識保持部106には、異なる複数の第1の変換知識が記憶されていてもよい。

#### 【0078】

図8に示した第1の変換知識は、推定結果記憶部104から、検索要求に含まれていたラベルと対応関係をもつ構成要素を検索するための第1の検索命令を生

成するためのもので、そのための記述が検索言語 XQUERY で記述されている。なお、ここで生成される第 1 の検索命令は、検索要求に含まれていたラベルと対応関係をもつ構成要素と、当該構成要素をもつ商品情報中の「店舗名」とを検索結果として求めるためのものである。第 1 の変換知識は、上記置換文字列以外は、予め定められた検索言語で記述されて、上記置換文字列を入力されたラベルに置き換えさえすれば第 1 の検索命令が完成するようになっている。

#### 【0079】

図 9 に示した第 2 の変換知識では、「##ROLE##」「##KEYWORD##」「##SHOP##」及び「##PATH##」と表現されている部分が置換文字列である。後述する手順でこれらの置換文字列が検索要求あるいは予備検索結果に従って置換され、検索言語 XQUERY で記述された第 2 の検索命令が生成される。なお、ここで生成される第 2 の検索命令は、検索要求に含まれていたラベルと対応関係にある構成要素を含むとともに、当該検索要求に含まれていたキーワードを要素値としてもつ構成要素を含む商品情報を検索し、この検索された商品情報の中から「店舗名」と、当該ラベルに対応する構成要素の要素値を検索結果として求めるためのものである。第 2 の変換知識は、上記置換文字列以外は、予め定められた検索言語で記述されて、上記置換文字列を、利用者により入力されたラベルやキーワード、予備検索の結果得られた店舗名やタグ名に置き換えさえすれば第 2 の検索命令が完成するようになっている。

#### 【0080】

本実施形態ではこれらの置換文字列は本装置での予約語として扱われるが、もちろん置換文字列の表現方法はここで示した例に限られない。例えばデータ中に決して現れないエスケープ文字によって置換文字列を表現すれば、予約語とデータの衝突を避けることができる。

#### 【0081】

また、図 9 では変換知識をテーブルによって表現しているが、例えば XML 表現によって記述して、構造化データ蓄積部 101 と同様に保持してもよい。

#### 【0082】

次に、図 10 に示すフローチャートを参照して検索要求変換部 107 の処理動

作を説明する。ここでは、第1の検索命令を生成するための処理動作について説明する。

#### 【0083】

ステップS21で、検索要求変換部107は検索要求入力部105から検索要求を受け取ると、続いてステップS22及びS23で、当該検索要求に含まれるラベルとキーワードをそれぞれ取得する。

#### 【0084】

次に、ステップS24において、第1の変換知識を、変換知識保持部106から取り出す。この取り出した変換知識の中に、置換文字列「##ROLE##」が含まれる場合（ステップS25）、これをステップS22で取得したラベルに置換える（ステップS26）。このステップS25からステップS26の処理を当該変換知識中の置換文字列「##ROLE##」が全て置換されるまで繰り返し（ステップS25）、その結果、第1の検索命令を生成する。

#### 【0085】

ステップS27では、当該第1の検索命令を検索部107へ出力する。検索部107では、当該第1の検索命令を基に、推定結果記憶部104から、当該第1の検索命令中に検索条件として定められたラベルに対応付けられた構成要素（すなわち、ここでは、タグ）と、当該タグを含む商品情報中の店舗名とを検索結果として求める、予備検索を開始する。なお、検索部107におけるXMLデータの検索手法は、公知、公用のXML検索システムなどと同様であり、また、本発明の要旨でないので、ここでは詳細を説明しない。

#### 【0086】

図11に示すフローチャートは、図10のステップS27で得られる、予備検索の結果を利用して、利用者の検索要求を第2の検索命令に変換するための検索要求変換部107の処理動作を説明するためのものである。

#### 【0087】

予備検索の結果として、タグ名と当該タグ名を含む商品情報中の店舗名とを1組の予備検索結果としたとき、少なくとも1組の予備検索結果が得られたとする。

。

## 【0088】

ステップS28では、予備検索結果を一組取り出し、当該予備検索結果から店舗名（ステップS29）とタグ（ステップS30）を取り出す。

## 【0089】

次に、図10のステップS22で取得したラベルに対応する第2の変換知識を、変換知識保持部106から取り出す（ステップS31）。取り出した第2の変換知識に置換文字列「##KEYWORD##」が含まれる場合（ステップS32）、これを図10のステップS23で取得した検索要求中のキーワードに置換し（ステップS33）、この処理を第2の変換知識中の置換文字列「##KEYWORD##」が全て置換されるまで繰り返す。

## 【0090】

また、取り出した第2の変換知識に置換文字列「##SHOP##」が含まれる場合（ステップS34）、これをステップS29で取得した、予備検索結果中の店舗名に置換し（ステップS35）、この処理を変換知識中の置換文字列「##SHOP##」が全て置換されるまで繰り返す。

## 【0091】

同様に、取り出した変換知識に置換文字列「##PATH##」が含まれる場合（ステップS36）、これをステップS30で取得した、予備検索結果中のタグ名に置換し（ステップS37）、この処理を変換知識中の置換文字列「##PATH##」が全て置換されるまで繰り返す。

## 【0092】

こうして、第2の検索命令が生成される。生成された第2の検索命令は検索部108に出力される（ステップS38）。

## 【0093】

予備検索結果が残っていれば、ステップS28に戻り、次の予備検索結果を取り出して上記同様に処理を繰り返し、予備検索結果が全て取り出されたならこのフローを終了する（ステップS39）。

## 【0094】

次に、例えば、図7（c）に示したようなラベルとキーワードを含む検索要求

がステップ S 2 1 で検索要求変換部 1 0 8 が受け取る場合を例にとり、図 1 0、図 1 1 に示した処理動作を具体的に説明する。

#### 【0095】

この場合、ステップ S 2 2 でラベルとして「価格」が取り出され、ステップ S 2 3 でキーワードとして「PC-B 2 0 0」が取り出される。ステップ S 2 4 では、図 8 に示したような第 1 の変換知識が取り出され、この第 1 の変換知識中の置換文字列「## R O L E ##」が探索され（ステップ S 2 5）、この置換文字列を、ステップ S 2 2 で取得されたラベル「価格」に置換する（ステップ S 2 6）。図 7 に示すように、この第 1 の変換知識には置換文字列「## R O L E ##」は一度しか出現しないので、ここではステップ S 2 6 は一度だけ処理される。

#### 【0096】

上記処理の結果、図 1 2 に示したような、第 1 の検索命令が生成される。この第 1 の検索命令は、検索部 1 0 8 へ渡されて、予備検索が開始される（ステップ S 2 7）。

#### 【0097】

第 1 の検索命令は、検索部 1 0 8 に送られ、予備検索処理が行なわれるが、この動作については既知の XML 検索システムなどと同様であり、ここでは詳細を説明しない。

#### 【0098】

この予備検索の結果として、例えば、図 1 4 に示すように、2 組（件）の予備検索結果が得られたとする。

#### 【0099】

次に、図 1 1 に処理動作を具体的に説明する。

#### 【0100】

ステップ S 2 8 で、図 1 4 に示した予備検索結果のうちの 1 つ目の予備検索結果「<タグリスト><店舗名> A A デンキ</店舗名><タグ>販売価格</タグ></タグリスト>」が取り出される。この予備検索結果から、ステップ S 2 9 では店舗名として「A A デンキ」、ステップ S 3 0 ではタグとして「販売価格」が取り出される。

## 【0101】

次に、ステップS22で検索要求から取り出されたラベル「価格」に対応する第2の変換知識として図9から第2の変換知識702が取り出される（ステップS31）。第2の変換知識702中の置換文字列「##KEYWORD##」はステップS23で検索要求から取り出されたキーワード「PC-B200」に置換される（ステップS32）。第2の変換知識702中に「##KEYWORD##」は一度しか出現しないのでステップS33は一度だけ処理される（ステップS32）。

## 【0102】

同様に第2の変換知識702中の置換文字列「##SHOP##」がステップS29で予備検索結果から取り出された店舗名「AAデンキ」に置換される（ステップS35）。第2の変換知識702中に「##KEYWORD##」は2回出現するので、ステップS35は2度処理される（ステップS34）。第2の変換知識702中の置換文字列「##PATH##」はステップS30で予備検索結果から取り出されたタグ「販売価格」に置換される（ステップS37）。第2の変換知識702中に「##PATH##」は一度しか出現しないので、ステップS37は一度だけ処理される（ステップS36）。

## 【0103】

このようにして生成されたXQUERY形式の第2の検索命令がステップS38で出力される。生成された第2の検索命令を図13に示す。

## 【0104】

なお、図14に示したように、予備検索結果は2件あるので、ステップS39によってステップS28から処理が繰り返される。

## 【0105】

2度目のループでは、ステップS28で予備検索結果として「<タグリスト><店舗名>YY商店</店舗名><タグ>TagC</タグ></タグリスト>」が取り出され、以降ステップS29～ステップS38が1度目のループと同様に処理される。

## 【0106】

予備検索結果は2件なので、2度目にステップS39に到達するとフローは終了する。

#### 【0107】

この例では検索要求変換部107によって、2つの第2の検索命令が出力されたことになる。1つ目の第2の検索命令に基づき、検索部108は、構造化データ蓄積部101から、店舗名「AAデンキ」の商品情報の中から、①検索要求に含まれていたキーワードを要素値としてもつ構成要素を含み、しかも②検索要求に含まれていたラベルと対応関係にある構成要素を含む、商品情報を検索する。また、2つ目の第2の検索命令に基づき「YY商店」の商品情報の中から、上記①および②を満たす商品情報を検索する。

#### 【0108】

上記1つ目の第2の検索命令によれば、検索された商品情報のうち、＜店舗名＞タグの値と、ラベル「価格」に対応する＜販売価格＞タグの値が検索結果として得られることになる。また、上記2つ目の第2の検索命令によれば、検索された商品情報のうち、＜店舗名＞タグの値と、ラベル「価格」に対応する＜Tag C＞タグの値が検索結果として得られることになる。

#### 【0109】

検索部108は、例えばXQUERYなどの検索言語を処理できる既知のXMLデータベース管理システムによって実現できるのでここでは動作の詳細を説明しない。

#### 【0110】

具体的な検索対象のXMLデータとして図2に示したような商品情報を考えると、図13に示した第2の検索命令によって、指定されたキーワードである「PC-B200」という文字列を含むデータとして図2(b)の商品情報が検索され、この検索された商品情報の内容が、図13に示した第2の検索命令に記述されている「RETURN」節によって再構成されて、検索結果として出力されることになる。

#### 【0111】

なお、図11の2度目のループで出力される上記2つ目の第2の検索命令につ

いては、図2に示した商品情報には、当該第2の検索命令中の上記検索条件①および②を満たす商品情報は存在しない。従って検索結果は得られない。

#### 【0112】

検索結果出力部109は、例えば、図15に示すように、上記検索結果を表示する。ここでは検索結果が1件のみの例を示しているが、複数の商品情報が検索された場合であっても、例えばリスト形式で表示すればよい。また、ここでは検索結果として検索部108から得られた表現をそのまま出力しているが、この場合に限らず、例えば「「PC-B200」は「AAデンキ」で「特価30%OFF」です」のように検索結果出力部109が自然文による表現などに変換して出力してもよい。

#### 【0113】

以上説明したように、上記第1の実施形態によれば、構造化データ蓄積部101に記憶された、複数の構成要素からなる階層化されたデータ構造をもつ複数の構造化データの構成要素の要素値の種類を推定して、当該構造化データの構成要素と、当該構成要素の要素値に対し推定された種類に対応するラベルとの対応関係を推定結果記憶部104に記憶しておく。キーワードとラベルを含む検索要求が入力されたとき、まず、当該検索要求に含まれるラベルに対応する構成要素を推定結果記憶部104に記憶されている対応関係から検索し、次に、この検索された構成要素と、当該検索要求に含まれるキーワードを要素値としてもつ構成要素とを含む構造化データを構造化データ蓄積部101に記憶された構造化データの中から検索する。この検索された構造化データのうち、少なくとも、検索要求に含まれるラベルに対応する構成要素の要素値を検索結果として出力する。

#### 【0114】

このような構成により、商品情報のデータ構造が未知であったり、従来の手段では要素値として含まれている個々のデータの意味役割といった要素値の種類がわからない場合であっても、キーワード（例えば製品名「PC-B200」）と要素値の種類に対応するラベル（例えば「価格」）を入力するだけで、当該キーワードを要素値としてもつ構成要素と、当該ラベルに対応する要素名をもつ構成要素とを含む所望の情報を検索し、当該キーワードを要素値としてもつ構成要素

を含む情報の中から、利用者の知りたい、当該ラベルに対応する要素名の構成要素の要素値（例えば「大特価 30%OFF」）を検索結果として得ることができる。

#### 【0115】

すなわち、異なるデータ構造の複数の構造化データの中から、所望の構造化データを検索する際には、そのデータ構造に依存することなく（正確な要素名を知る必要なく、検索したいデータの種類に対応するラベルを指定すれば）、所望の情報データを要素値としてもつ構成要素を含む構造化データを容易に、しかも確実に検索することができる。

#### 【0116】

（第1実施形態の第1の変形例）

次に上記第1実施形態の第1の変形例を説明する。この第1の変形例では、ユーザからの自然言語による検索要求の入力を受け付けることが可能で、入力された自然言語による検索要求からラベルを推定し、この推定結果に基づいた検索を行なうことができる、商品情報検索装置を考える。

#### 【0117】

図16は、第1の変形例に係る商品情報検索装置の構成例を示したものである。なお、図16において、図1と同一部分には同一符号を付し、異なる部分についてのみ説明する。すなわち、図16では、検索要求入力部105からユーザにより入力される自然文の検索要求から、ラベルやキーワードに対応する語を推定するための、第2の推定部111と第2の推定知識保持部110と、さらに、第2の推定部111で推定されたラベルやキーワードに対応する語、さらに、これらを基に生成された第1の検索命令や第2の検索命令をユーザに提示して、必要に応じて修正させるための修正部112が追加されている。

#### 【0118】

検索要求入力部105は、ユーザから入力される自然言語で記述された検索要求を受け付け、それを検索要求変換部107に送る。

#### 【0119】

第1の変形例では検索言語として第1実施形態と同様XQUERYを考える。

## 【0120】

第2の推定知識保持部110は、検索要求入力部105から入力された検索要求からラベルを推定するための推定知識を保持する。

## 【0121】

検索要求変換部107は、検索要求入力部105から渡された検索要求を、まず、第2の推定部111へ送り、ここで、第2の推定知識保持部110に記憶されている推定知識に基づき、当該自然文の検索要求から、第1の実施形態で説明した、ラベルや、キーワードに対応する語を推定する。その後、検索要求変換部107では、この推定されたラベルや、キーワードに対応する語を基に、第1の実施形態で説明したように、第1の検索命令や第2の検索命令を生成する。生成された第1の検索命令や第2の検索命令は、修正部112に送られる。

## 【0122】

修正部112は、第2の推定部111で推定されたラベルやキーワードに対応する語、検索要求変換部107によって生成された第1の検索命令、第2の検索命令をユーザに提示し、ユーザからの修正を受け付け、修正されたラベルやキーワードを再び、検索要求変換部107へ渡したり、修正された第1の検索命令や第2の検索命令を検索部108に送る。もちろん、ラベルやキーワード、第1の検索命令や第2の検索命令に修正がなければ、そのまま検索要求変換部107、検索部108に送られる。

## 【0123】

以下、上記各部に処理動作について具体的に説明する。

## 【0124】

図17は、第1の変形例における検索要求入力部105により表示される検索要求入力画面の一例を示したものである。ここでは検索要求として「DB3254はいくらですか?」という自然文が入力されている。

## 【0125】

検索要求変換部107は、検索要求入力部105から上記検索要求を受け取り、これを第2の推定部111へと送ることにより、上記検索要求からラベルと、キーワードに対応する語を推定結果として受け取る。

## 【0126】

第2の推定知識保持部110には、例えば図18に示すような推定知識が記憶されている。ここに記憶されている推定知識は、ラベルと、検索要求の自然文に含まれることが予想される当該ラベル（すなわち、構成要素の要素値の種類を表すような語）を指定していると推定される語とを対応付けて構成されたものである。ここでは、構成要素の要素値の種類を表すような語を指定していると推定される語をパターンと呼ぶ。図18に示す推定知識によれば、例えば、検索要求の自然文中に、「値段」、「価格」、「いくら」などといった語が含まれていれば、このような語により、当該検索要求により「価格」というラベルが指定されているものと推定する。

## 【0127】

次に、図19に示すフローチャートを参照して第2の推定部111におけるラベルと、キーワードに対応する語の推定処理動作について説明する。

## 【0128】

第2の推定部111は、検索要求入力部105から入力された検索要求を受け取ると、当該検索要求を形態素解析して、当該検索要求から単語を抽出する（ステップS41）。第2の推定知識保持部110に記憶されている図18に示したような推定知識を1つずつ取り出し、当該抽出した単語の中に、各推定知識のパターンに一致するものがあるかチェックする（ステップS42）。推定知識のパターンと一致する単語があれば（ステップS43）、当該推定知識のラベルを当該検索要求のラベルと推定するとともに、ステップS41で抽出された単語のうち、当該ラベルに対応する語以外の単語のうち自立語があれば、それを、キーワードとして抽出する（ステップS44）。一方、ステップS43で、推定知識のパターンと一致する単語がなければ、ラベルは「不定」と推定し、ステップS41で抽出された単語のうち自立語があれば、それをキーワードとして抽出する（ステップS45）。

## 【0129】

例えば、図17に示したような自然文の検索要求が入力された場合、当該自然文から、「いくら」という語がラベル「価格」のパターンに一致するので、ラベル

として「価格」が求まる。また、当該自然文中の単語のうち、「いくら」以外の単語からは、「DB3254」が自立語であるので、この語が、キーワードとなる。

#### 【0130】

第2の推定部111での推定結果は、第1の実施形態で検索要求入力部105に入力された検索要求と同様に「ラベル＝“価格”、キーワード＝“DB3254”」といった表現で保持することができる。

#### 【0131】

このように推定結果を保持すれば、第1の変形例の検索要求変換部107による変換処理は第1の実施形態の場合と同様、図10、図11に従って行なうことができる。

#### 【0132】

ただし、第1の変形例では、第2の推定部111において、図19のステップS45のように、検索要求からのラベルの推定に失敗する場合がある（ラベルが「不定」と推定される場合）。そこで、このラベルが「不定」の場合の第2の変換知識を図20に示す。

#### 【0133】

図20に示す、ラベルが「不定」の場合の第2の変換知識では、図19のステップS45でキーワードとして抽出された語を要素値としてもつ構成要素を含む商品情報を検索するための第2の検索命令が生成されるようになっている。このようなラベルが「不定」の第2の変換知識も変換知識保持部106に記憶する。

#### 【0134】

第1の変形例での検索要求変換部107で生成される第2の変換命令のうちの1つを図21に示す。第1実施形態同様、検索要求変換部107によって生成される第2の検索命令は複数出力されることがある。

#### 【0135】

修正部112は、図22に示すような修正画面を表示する。この修正画面は、ユーザにより入力された自然文の検索要求を表示（および修正）する領域801と、第2の推定部111において、当該検索要求から抽出されたキーワードに対

応する語を表示（および修正）する領域 802 と、当該検索要求から推定されたラベルを表示（および修正）する領域 803 とからなる。

#### 【0136】

ユーザは、この修正画面の領域 802、803 に表示されたキーワード、ラベルに修正があれば、各領域 802、803 上で修正することができる。

#### 【0137】

図 22 に示す修正画面上で、ユーザは第 2 の推定部 111 の推定結果を直接修正することもできるし、自然文の検索要求自体を再度入力し直して修正作業を行なってもよい。

#### 【0138】

領域 801 に検索要求の入力をやり直した場合には、第 2 の推定部 111 で再度図 19 に示した推定処理が行われ、その推定結果が再び、図 22 の領域 802、803 に表示される。

#### 【0139】

ユーザは必要があれば検索要求の修正を行ない、所望の検索条件が設定されたら、検索の実行を指示するためのボタン 804 を押す。

#### 【0140】

ここでは特にユーザからの検索要求の修正はなかったものとする。検索要求の修正が完了すると、修正部 112 は、キーワード、ラベルを検索要求変換部 107 へ渡し、ここで、第 1 の実施形態で説明したように、第 1 の検索命令、第 2 の検索命令が生成される。以後の処理動作は、第 1 の実施形態と同様である。

#### 【0141】

以上説明したように、上記第 1 の変形例によれば、構造化データ蓄積部 101 に記憶された、複数の構成要素からなる階層化されたデータ構造をもつ複数の構造化データの構成要素の要素値の種類を推定して、当該構造化データの構成要素と、当該構成要素の要素値に対し推定された種類に対応するラベルとの対応関係を推定結果記憶部 104 に記憶しておく。自然文の検索要求が入力されたとき、当該検索要求としての自然文に含まれる語から前記ラベルを推定するとともに、キーワードに対応する語を推定する。そして、まず、この推定されたラベルに対

応する構成要素を推定結果記憶部 104 に記憶された対応関係から検索し、次に、検索された構成要素と、当該推定されたキーワードを要素値としてもつ構成要素とを含む構造化データを構造化データ蓄積部 101 に記憶された構造化データの中から検索する。この検索された構造化データのうち、少なくとも、検索要求から推定されたラベルに対応する構成要素の要素値を検索結果として出力する。

#### 【0142】

このような構成により、商品情報のデータ構造が未知であったり、従来の手段では要素値として含まれている個々のデータの意味役割といった要素値の種類がわからない場合であっても、ユーザが質問を自然言語によって記述すると（例えば「DB3254はいくらですか？」を入力すると）、その質問からラベルとキーワードを推定することにより、当該キーワードを要素値としてもつ構成要素と、当該ラベルに対応する要素名をもつ構成要素とを含む所望の情報を検索し、当該キーワードを要素値としてもつ構成要素を含む情報の中から、利用者の知りたい、当該ラベルに対応する要素名の構成要素の要素値（例えば「キャンペーン価格」）を回答として得ることができる。

#### 【0143】

（第1実施形態の第2の変形例）

第1実施形態の第1の変形例のさらに変形例として、修正部 112 が、検索要求変換部 107 で生成された第1の検索命令や第2の検索命令をユーザに提示し、修正を受け付ける場合を考える。

#### 【0144】

この第2の変形例に係る商品情報検索装置の構成例を図23に示す。なお、図23において、図1や図16と同一部分には同一符号を付し、異なる部分についてのみ説明する。すなわち、図23に示す構成は、図16に示した構成にさらに、第1の推定部 103 での推定結果をユーザに表示するための推定結果提示部 113 が追加されている。

#### 【0145】

例えば、図2に示したような商品情報が構造化データ蓄積部 101 に記憶されており、また、第2の推定知識保持部 110 には図18に示したような推定知識

が記憶され、変換知識保持部 106 には、図 8、図 9 に示したような変換知識が記憶されているとする。このとき、ユーザにより検索要求として「DB 3254 はいくらですか?」が入力された場合を考える。

#### 【0146】

この場合、上記検索要求からは、前述したように、ラベルとして「価格」が推定され、キーワードとして「DB 3254」が推定されるので、これらから前述したように、第 1 の検索命令、第 2 の検索命令が生成される。

#### 【0147】

例えば、第 2 の検索命令が生成されると、修正部 112 は、当該第 2 の検索命令の修正を可能にするために、当該第 2 の検索命令を表示し、推定結果提示部 113 は、推定結果記憶部 104 に記憶されている推定結果を表示する。

#### 【0148】

図 24 は、上記第 2 の検索命令の修正画面の表示例を示したもので、この修正画面には、上記第 2 の検索命令と推定結果とが表示されている。図 24 に示したように、第 2 の検索命令と推定結果とは同一画面上に表示され、第 2 の検索命令は、領域 902 に表示され、検索結果は領域 901 に表示される。

#### 【0149】

ユーザは、図 24 に示した画面上で、領域 901 に提示された推定結果を参照しながら、領域 902 に表示された検索要求変換部 107 で生成した第 2 の検索命令を修正することができるようになっている。ユーザは、例えば、図 24 に示した画面上で、第 2 の検索命令に「<時間> { \$ a / データ / 稼働時間 / t e x t ( ) } < / 時間 > 」という 1 行を追加している。

#### 【0150】

この修正により、ユーザは予め入力した検索要求「DB 3254 はいくらですか?」に加え、時間に関する情報、恐らくはバッテリーの持続時間についての情報も検索するように指示している。

#### 【0151】

第 2 の検索命令の修正が済むと、ユーザは実行ボタン 904 を押す。すると、修正部 112 は、当該修正された第 2 の検索命令を検索部 108 に送る。その後

の検索部108、検索結果出力部109の処理動作は、第1実施形態と同様なのでここでは説明しない。

#### 【0152】

第2の変形例の場合、ユーザは商品情報のデータ構造を参照しながら、より詳細な検索の指示が行なえることになる。

#### 【0153】

また、ユーザはある程度検索言語（例えばXQuery）に関する知識を持っている必要があるが、初期の検索要求を自然言語で与えることができるので、最初から検索言語を用いて検索命令を構築するよりも簡単に検索を行なうことができる。

#### 【0154】

本発明の実施の形態に記載した本発明の手法は、コンピュータに実行させることのできるプログラムとして、磁気ディスク（フロッピーディスク、ハードディスクなど）、光ディスク（CD-ROM、DVDなど）、半導体メモリなどの記録媒体に格納して頒布することもできる。

#### 【0155】

なお、本発明は、上記実施形態に限定されるものではなく、実施段階ではその要旨を逸脱しない範囲で種々に変形することが可能である。さらに、上記実施形態には種々の段階の発明は含まれており、開示される複数の構成要件における適宜な組み合わせにより、種々の発明が抽出され得る。例えば、実施形態に示される全構成要件から幾つかの構成要件が削除されても、発明が解決しようとする課題の欄で述べた課題（の少なくとも1つ）が解決でき、発明の効果の欄で述べられている効果（の少なくとも1つ）が得られる場合には、この構成要件が削除された構成が発明として抽出され得る。

#### 【0156】

##### 【発明の効果】

以上説明したように、本発明によれば、異なるデータ構造の複数の構造化データの中から、所望の構造化データを検索する際には、そのデータ構造に依存することなく、所望の情報データを要素値としてもつ構成要素を含む構造化データを

容易に、しかも確実に検索することができる。

【図面の簡単な説明】

【図 1】

本発明の第 1 の実施形態に係る商品情報検索装置の構成例を示した図。

【図 2】

商品情報の具体例を示した図。

【図 3】

第 1 の推定知識保持部に保持されている推定知識の具体例を示した図。

【図 4】

第 1 の推定部の処理動作を説明するためのフローチャート。

【図 5】

推定結果記憶部に記憶されているラベルとタグとの対応関係の記憶例を示した図。

【図 6】

推定結果記憶部に記憶されているラベルとタグとの対応関係の他の記憶例を示した図。

【図 7】

検索要求を入力するための手順を説明するための図。

【図 8】

変換知識保持部に記憶されている第 1 の変換知識の一例を示した図。

【図 9】

変換知識保持部に記憶されている第 2 の変換知識の例を示した図。

【図 10】

検索要求変換部の処理動作のうち、第 1 の検索命令を生成するための処理動作を説明するためのフローチャート。

【図 11】

検索要求変換部の処理動作のうち、第 2 の検索命令を生成するための処理動作を説明するためのフローチャート。

【図 12】

第1の検索命令の具体例を示した図。

【図13】

第2の検索命令の具体例を示した図。

【図14】

第1の検索命令を実行して（予備検索を行って）得られた検索結果の一例を示した図。

【図15】

第2の検索命令を実行した結果である検索結果の表示例を示した図。

【図16】

第1の実施形態の第1の変形例に係る商品情報検索装置の構成例を示した図。

【図17】

自然文の検索要求の入力画面の一例とその入力画面に入力された検索要求を示した図。

【図18】

第2の推定知識保持部に記憶されている推定知識の具体例を示した図。

【図19】

第2の推定部におけるラベルと、キーワードに対応する語の推定処理動作を説明するためのフローチャート。

【図20】

ラベルが「不定」の場合に用いる第2の変換知識の例を示した図。

【図21】

検索要求変換部で生成される第2の変換命令の一例を示した図。

【図22】

入力された検索要求から推定されたラベルとキーワードを修正する修正画面の一例を示した図。

【図23】

第1の実施形態の第2の変形例に係る商品情報検索装置の構成例を示した図。

【図24】

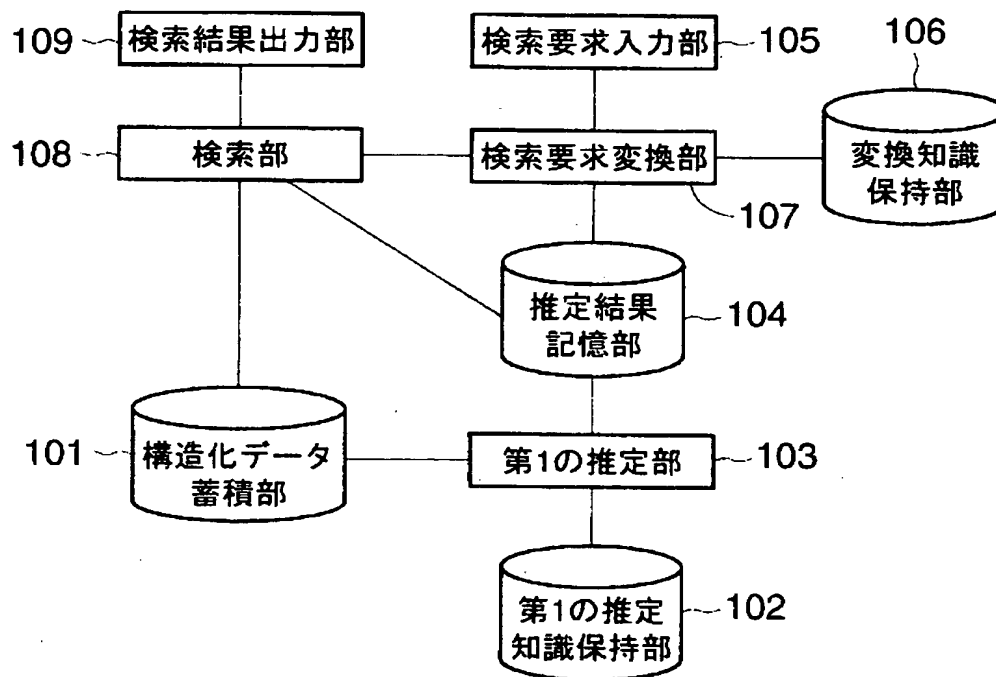
第2の検索命令の修正画面の表示例を示した図。

## 【符号の説明】

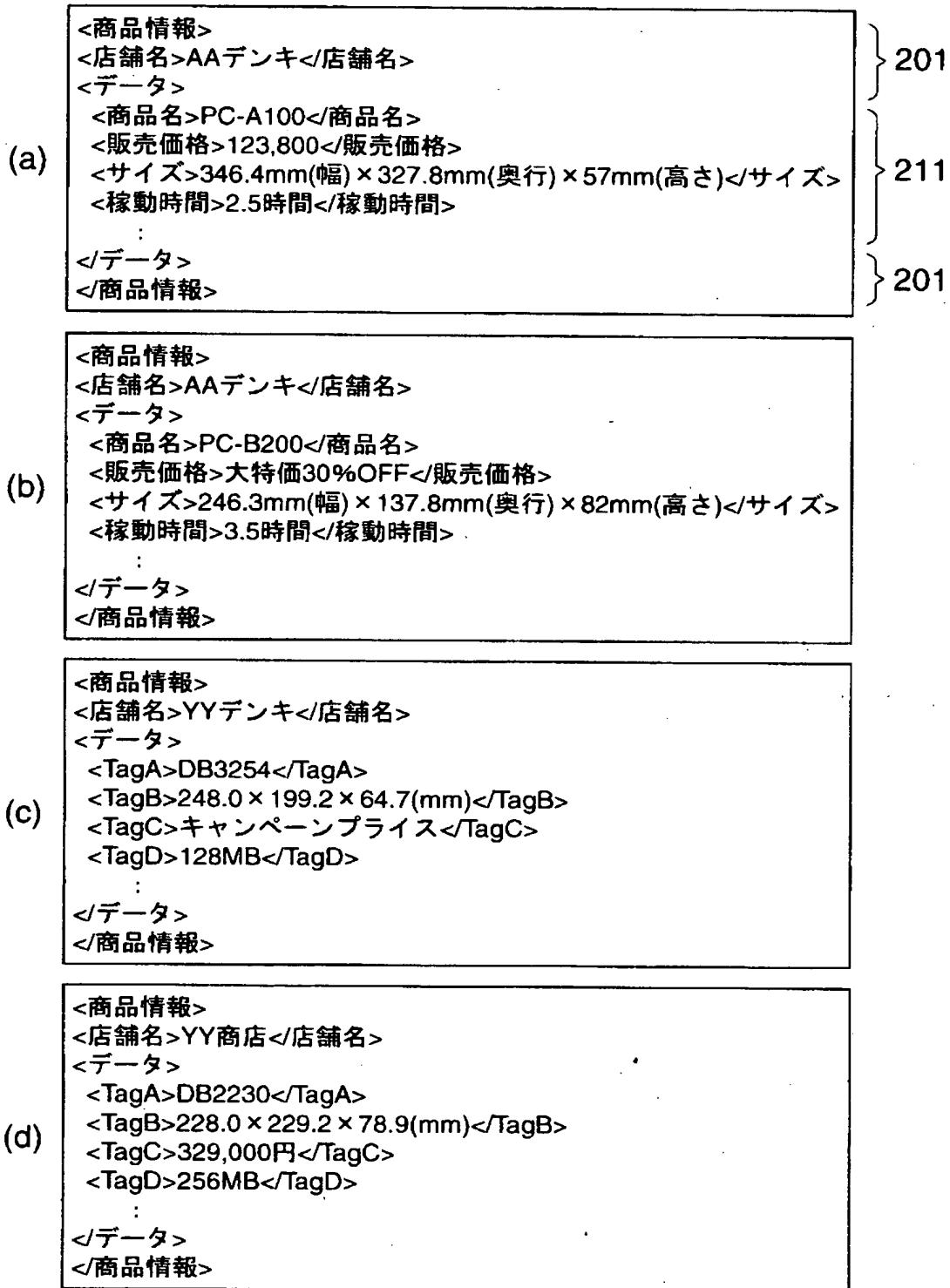
- 1 0 1 … 構造化データ蓄積部
- 1 0 2 … 第 1 の推定知識保持部
- 1 0 3 … 第 1 の推定部
- 1 0 4 … 推定結果記憶部
- 1 0 5 … 検索要求入力部
- 1 0 6 … 変換知識保持部
- 1 0 7 … 検索要求変換部
- 1 0 8 … 検索部
- 1 0 9 … 検索結果出力部
- 1 1 0 … 第 2 の推定知識保持部
- 1 1 1 … 第 1 の推定部
- 1 1 2 … 修正部
- 1 1 3 … 推定結果提示部

【書類名】 図面

【図 1】



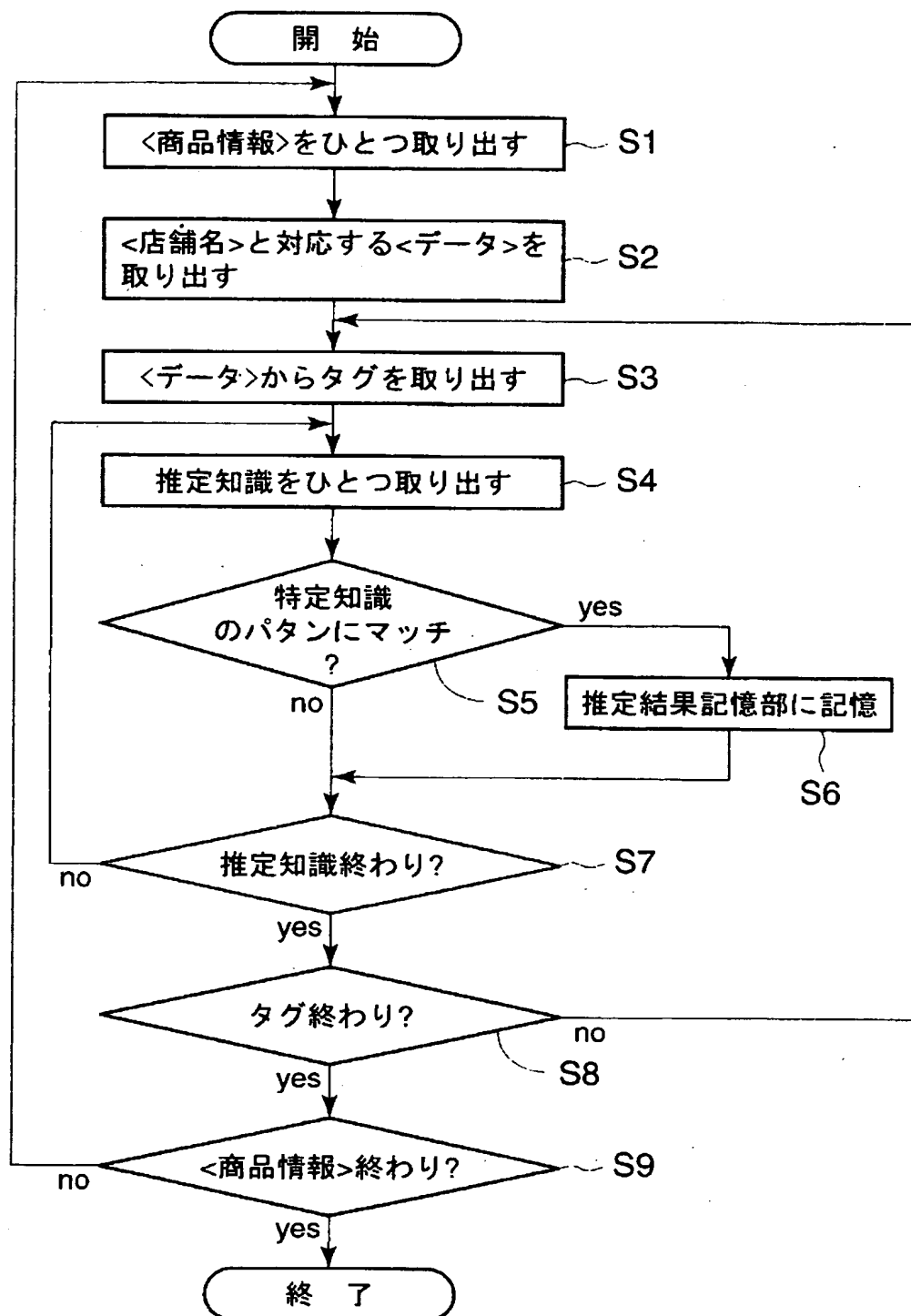
【図 2】



【図 3】

	パターン	ラベル
301	[¥d,]+円	価格
302	¥ ¥[¥d,]+	価格
303	[¥d.]+時間	時間
304	¥ d+時間 ¥ d+分	時間
305	[¥d.]+[cm]?m	長さ
306	¥ d+[GM]B	容量
307	[¥d.]+[GM]Hz	周波数
	:	:

【図 4】



【図5】

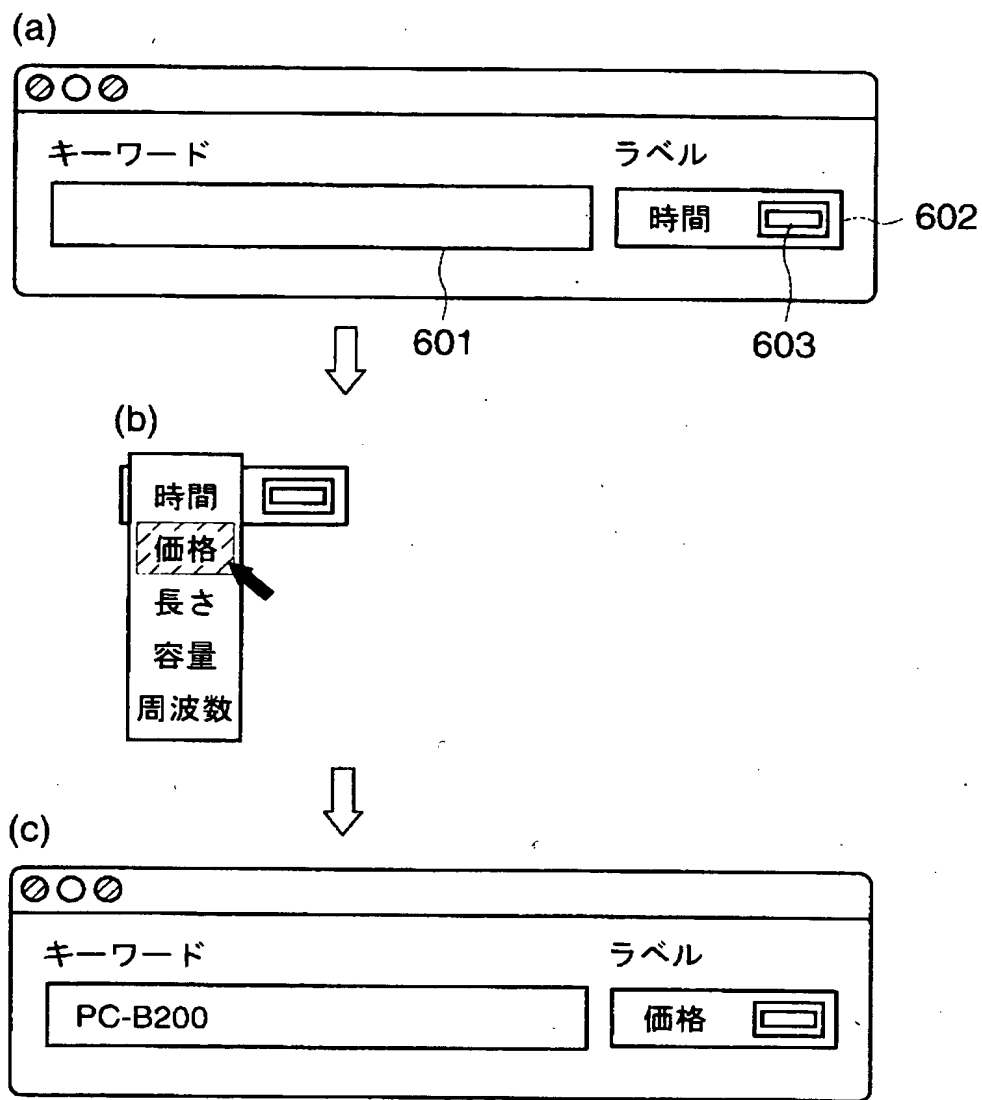
501 ~

店舗名	タグ	ラベル
AAデンキ	販売価格	[価格]
AAデンキ	サイズ	[長さ]
AAデンキ	稼動時間	[時間]
ABCストア	Memory	[容量]
ABCストア	CPU	[周波数]
YY商店	TagB	[長さ]
YY商店	TagD	[容量]
YY商店	TagC	[価格]
:	:	:

【図6】

<推定結果>  
 <意味役割ラベル>価格</意味役割ラベル>  
 <店舗名>AAデンキ</店舗名>  
 <タグ>販売価格</タグ>  
 </推定結果>

【図 7】



【図 8】

## 第1の変換知識

```
FOR $a IN/推定結果[ラベル=" ##ROLE##" ]
RETURN <タグリスト>{$a/店舗名}{$a/タグ}</タグリスト>
```

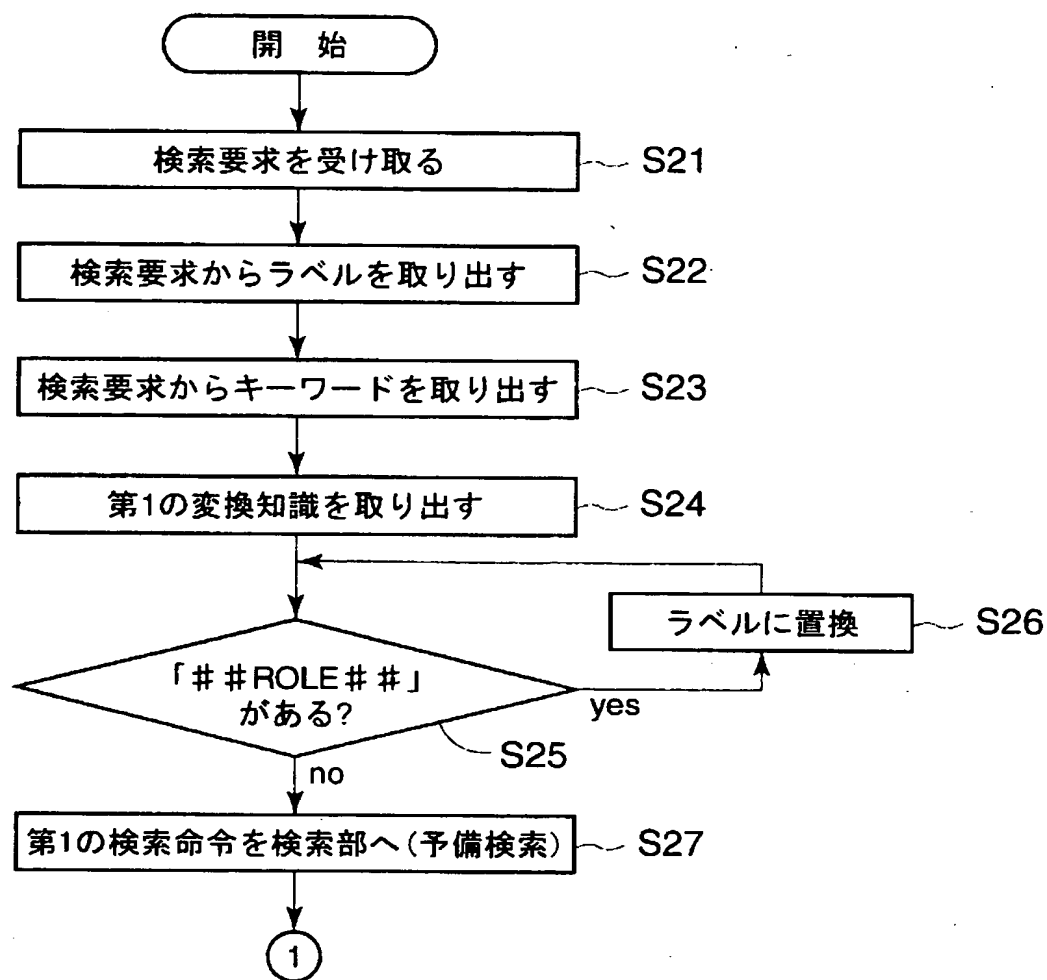
【図 9】

## 第2の変換知識

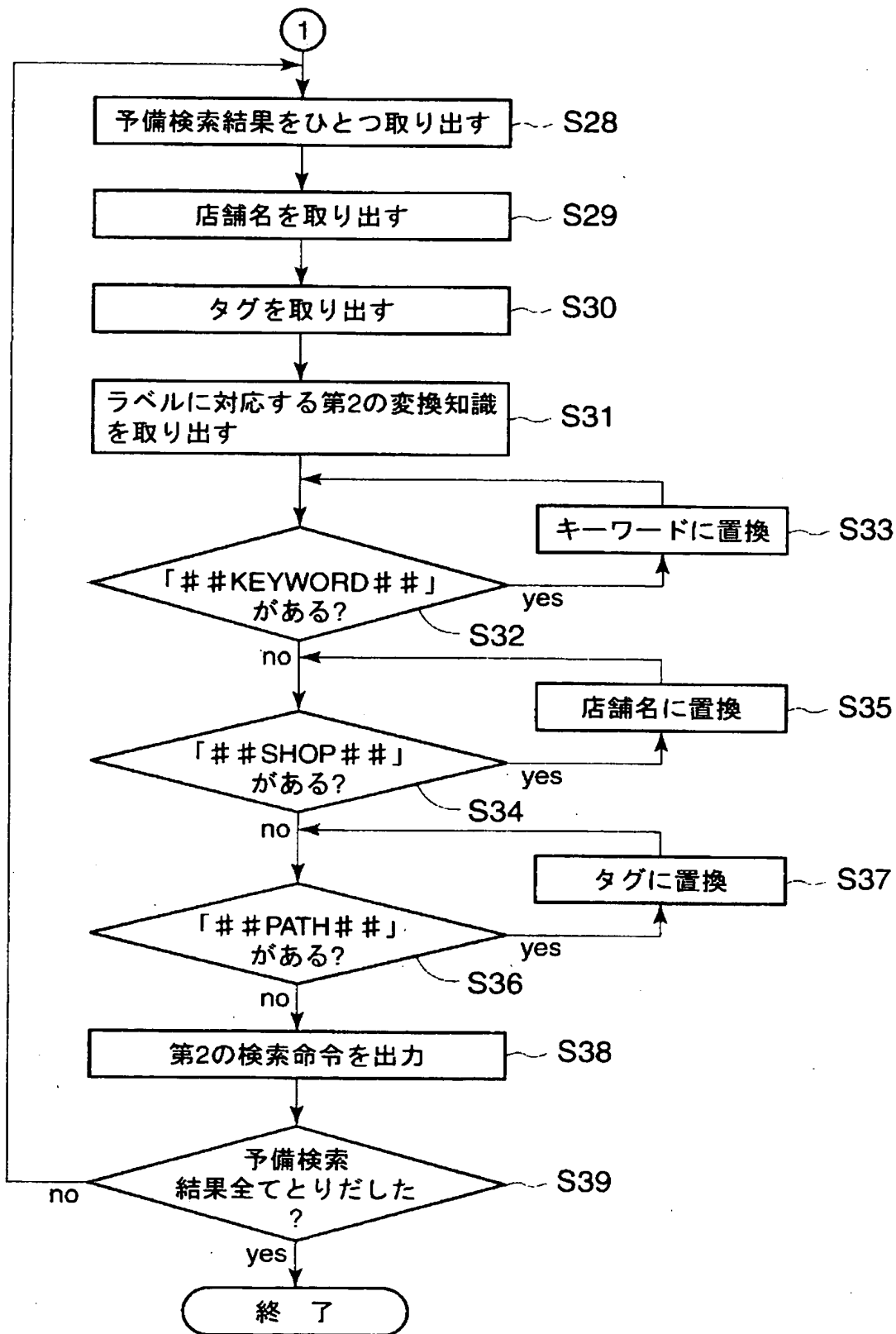
意味役割 ラベル	変換知識
[価格]	<pre>FOR \$a IN/商品情報 WHERE contains{\$a/text()," ##KEYWORD##" }AND \$a/店舗名= " ##SHOP##" RETURN &lt;商品情報&gt; &lt;店舗名&gt;##SHOP##&lt;/店舗名&gt; &lt;価格&gt;{\$a/データ/##PATH##/text()}&lt;価格&gt; &lt;/商品情報&gt;</pre>
[時間]	<pre>FOR \$a IN/商品情報 WHERE contains{\$a/text()," ##KEYWORD##" }AND \$a/店舗名= " ##SHOP##" RETURN &lt;商品情報&gt; &lt;店舗名&gt;##SHOP##&lt;/店舗名&gt; &lt;時間&gt;{\$a/データ/##PATH##/text()}&lt;時間&gt; &lt;/商品情報&gt;</pre>
⋮	⋮

702

【図10】



【図11】



## 【図 12】

## 第1の検索命令

```
FOR $a IN/推定結果[ラベル=" 価格" ]  
RETURN <タグリスト>{$a/店舗名}{$a/タグ}</タグリスト>
```

## 【図 13】

## 第2の命令

```
FOR $a IN/商品情報  
WHERE contains($a/text()," PC-B200" )AND $a/店舗名=" AAデンキ"  
RETURN  
<商品情報>  
<店舗名>AAデンキ</店舗名>  
<価格>{$a/データ/販売価格/text()}</価格>  
</商品情報>
```

## 【図 14】

## 予備検索の結果

```
<タグリスト><店舗名>AAデンキ</店舗名><タグ>販売価格</タグ></タグリスト>  
<タグリスト><店舗名>YY商店</店舗名><タグ>TagC</タグ></タグリスト>
```

【図15】

The image shows a web application window with a title bar containing three small circles. The main content area is divided into two sections. The top section has two labels: 'キーワード' (Keyword) and 'ラベル' (Label). Below 'キーワード' is a text input field containing 'PC-B200'. Below 'ラベル' is a button labeled '価格' (Price) with a small icon of a price tag. The bottom section is titled '検索結果' (Search Results) and contains a large rectangular box with the following text: '<商品情報>' (Product Information), '<店舗名>AAデンキ</店舗名>' (Store Name: AA Denki), '<価格>大特価30%OFF</価格>' (Price: Special Price 30% OFF), and '</商品情報>' (End of Product Information).

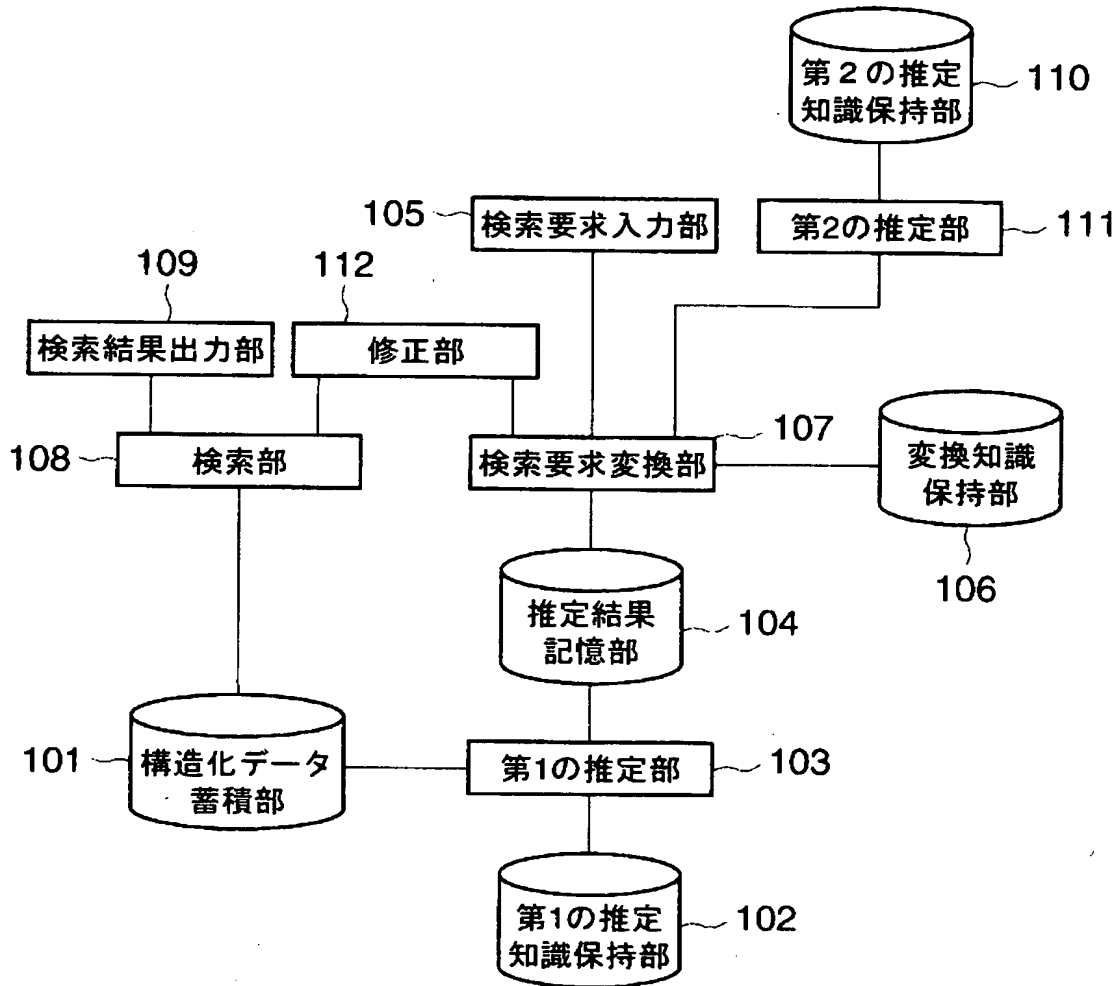
キーワード	ラベル
PC-B200	価格

---

検索結果

<商品情報>  
<店舗名>AAデンキ</店舗名>  
<価格>大特価30%OFF</価格>  
</商品情報>

【図 16】



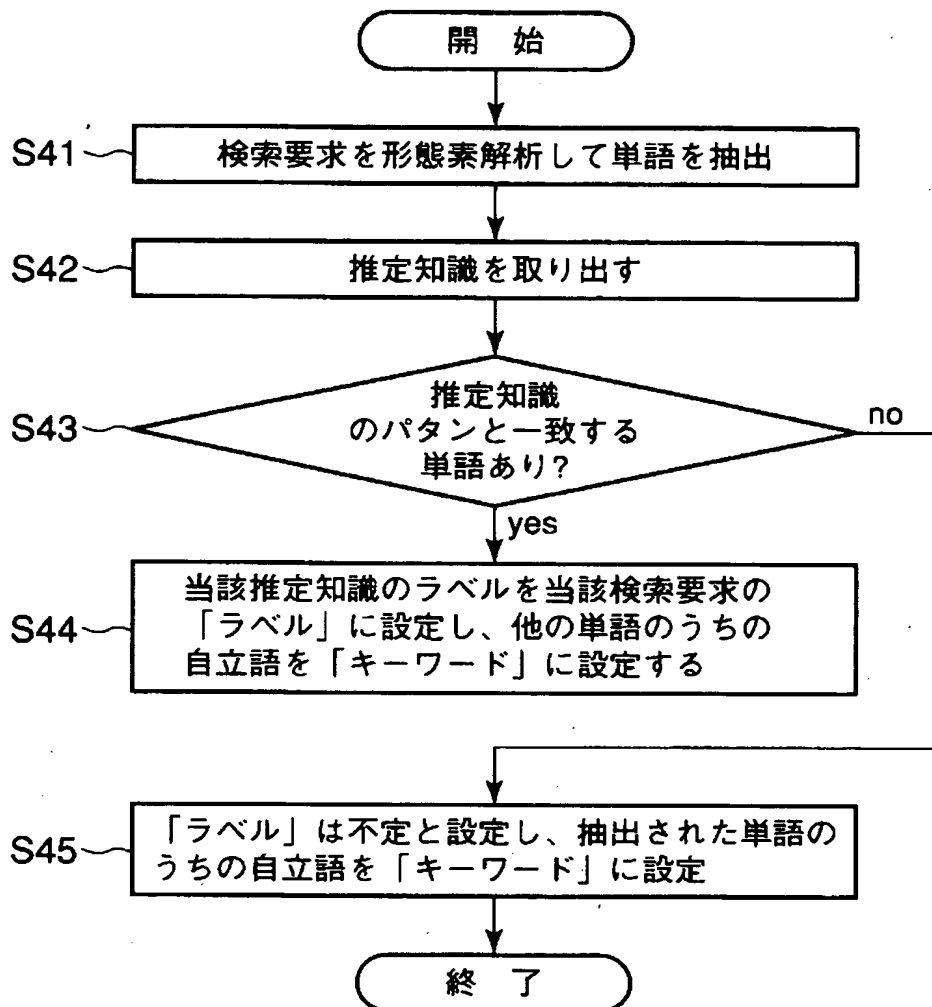
【図 17】

The screenshot shows a window with a title bar containing three circles. The main area contains the text "ご質問は?" (What is your question?). Below this text is a text input field containing the text "DB3254はいくらですか?" (How much is DB3254?).

【図 18】

パターン	ラベル
値段<名>	[価格]
価格<名>	[価格]
いくら<名>	[価格]
サイズ<助数>	[長さ]
時間<名>	[時間]
CPU<英>	[周波数]
:	:

【図 19】



【図 20】

ラベル	変換知識
[不定]	FOR \$a IN/商品情報 WHERE contains(\$a/text()," ##KEYWORD##" )AND \$a/店舗名= " ##SHOP##" RETURN {\$a}

【図 21】

```

FOR $a IN/商品情報
WHERE contains($a/text()," DB3254" )AND $a/店舗名=" AAデンキ"
RETURN
<商品情報>
<店舗名>AAデンキ</店舗名>
<価格>{$a/データ/販売価格/text()}</価格>
</商品情報>

```

【図 22】

ご質問は?

DB3254はいくらですか?

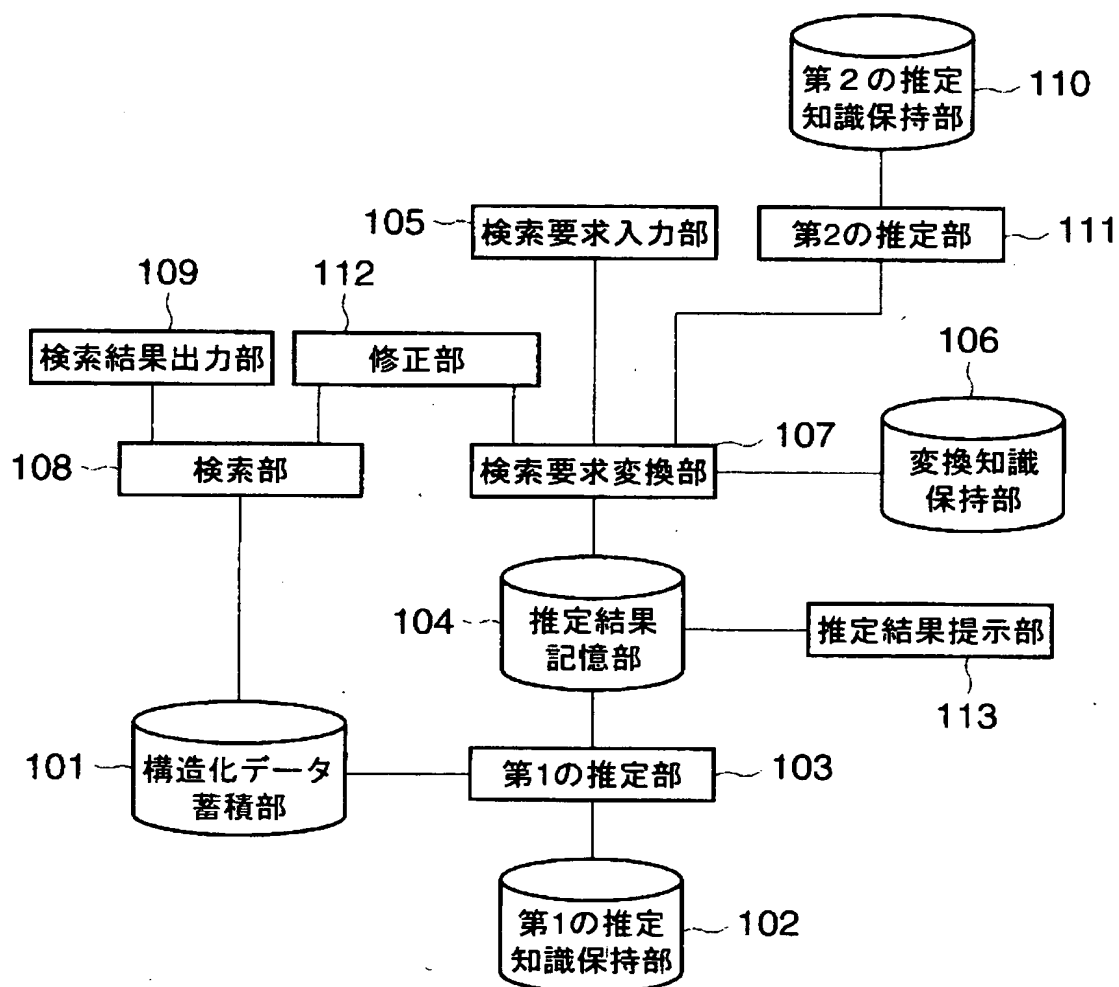
必要があれば修正して実行ボタンを押してください。

キーワード                      ラベル

DB3254                      価格

実行

【図 23】



【図 24】

901

データ構造テーブル

店舗名	タグ	ラベル
AAデンキ	販売価格	[価格]
AAデンキ	サイズ	[長さ]
AAデンキ	稼動時間	[時間]
ABCストア	Memory	[容量]
ABCストア	CPU	[周波数]
YY商店	TagB	[長さ]
YY商店	TagD	[容量]
YY商店	TagC	[価格]
:	:	:

検索命令 (必要があれば修正して実行ボタンを押してください)

```

FOR $a IN/商品情報
WHERE contains( $a/text(), " DB3254" )AND $a/店舗名=" AAデンキ"
RETURN
<商品情報>
<店舗名>AAデンキ</店舗名>
<価格>{ $a/データ/販売価格/text() }</価格>
<時間>{ $a/データ/稼動時間/text() }</時間>
</商品情報>

```

903

実行

902

904

【書類名】 要約書

【要約】

【課題】異なるデータ構造の複数の構造化データの中から、所望の構造化データを検索する際には、そのデータ構造に依存することなく（正確な要素名を知る必要なく）、所望の情報データを要素値としてもつ構成要素を含む構造化データを容易に、しかも確実に検索することができる構造化データ検索装置を提供する。

【解決手段】記憶手段に記憶された複数の構造化データの構成要素の要素値の種類を推定して、当該構造化データの構成要素と、当該構成要素の要素値に対し推定された種類に対応するラベルとの対応関係を記憶しておく。キーワードとラベルを含む検索要求が入力されたとき、まず、当該検索要求に含まれるラベルに対応する構成要素を上記対応関係から検索し、次に、この検索された構成要素と、当該検索要求に含まれるキーワードを要素値としてもつ構成要素とを含む構造化データを記憶手段に記憶された構造化データの中から検索する。

【選択図】 図1

特願 2002-327127

出願人履歴情報

識別番号

[000003078]

1. 変更年月日            2001年   7月   2日  
    [変更理由]          住所変更  
                          住 所      東京都港区芝浦一丁目1番1号  
                          氏 名      株式会社東芝
  
2. 変更年月日            2003年   5月   9日  
    [変更理由]          名称変更  
                          住所変更  
                          住 所      東京都港区芝浦一丁目1番1号  
                          氏 名      株式会社東芝